Kavli Institute for Systems Neuroscience
Centre for Neural Computation

NORDITA

Spin glasses and Computation

Yasser Roudi

Kavli Institute & Centre for Neural Computation
NTNU, Trondheim

Nordita, Stockholm

NETADIS
Statistical Physics Approaches to Networks Across Disciplines

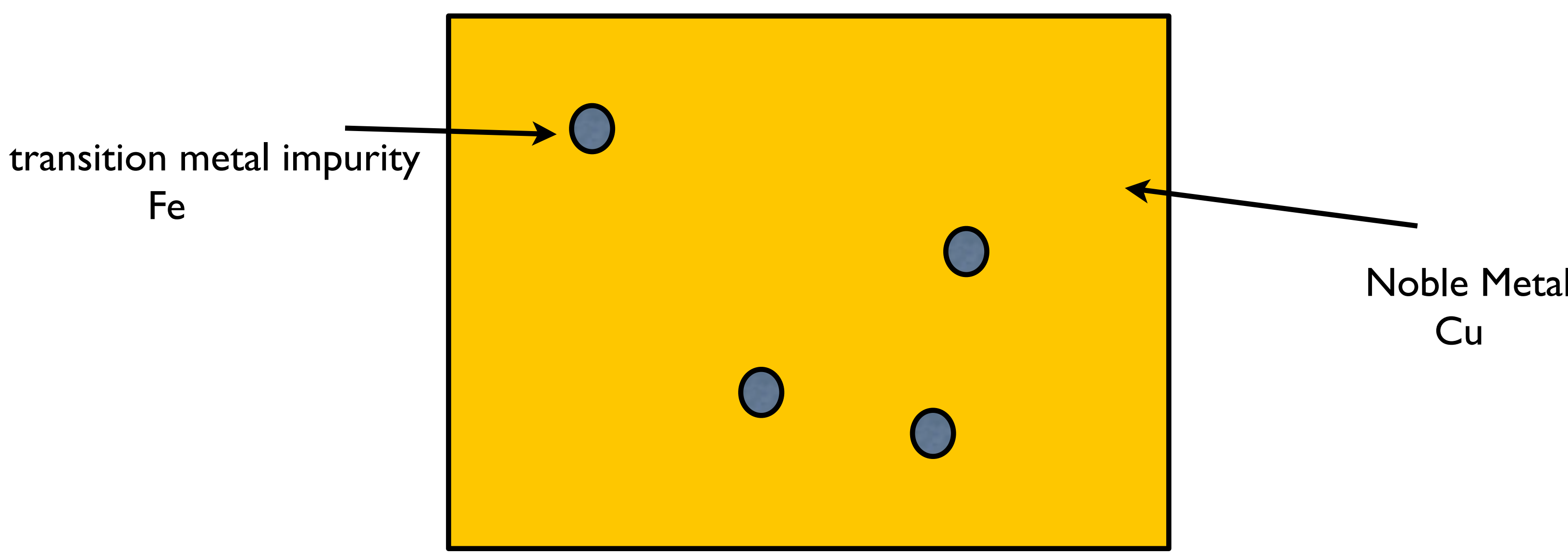
outline

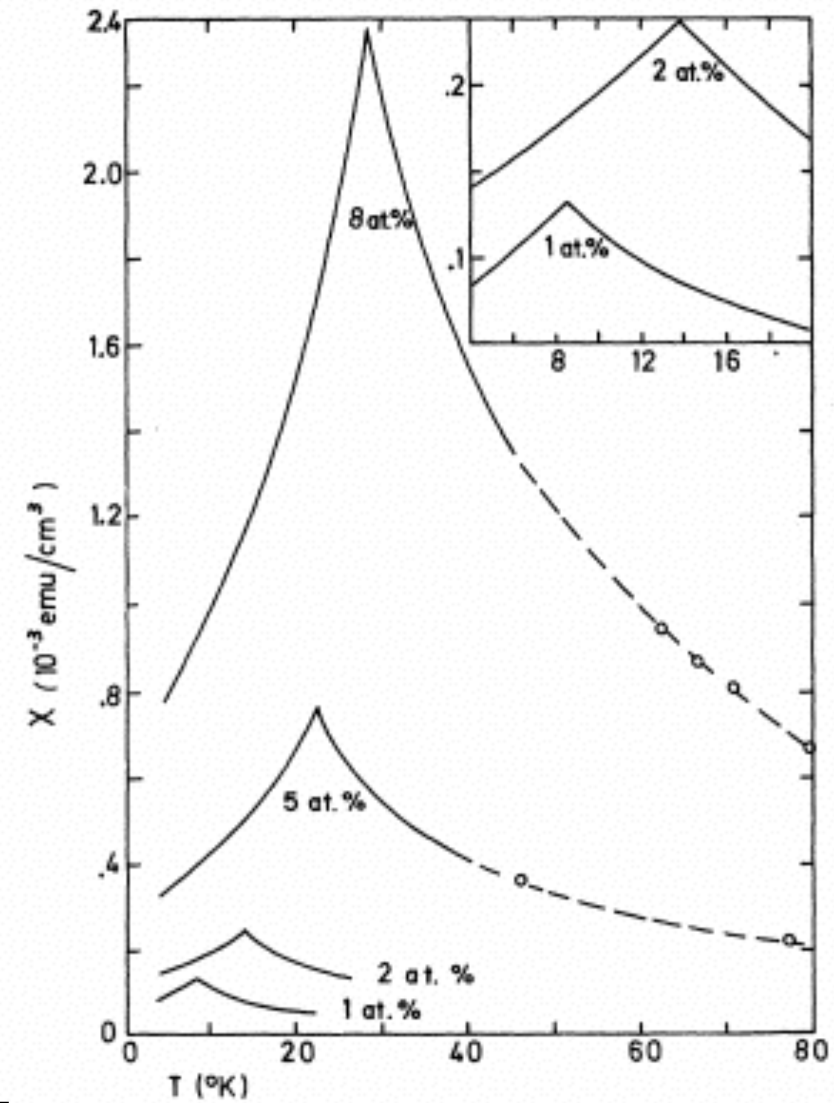
- What is a spin glass?
- how do we study spin glasses?
- how can we use spin glass physics to study information processing?
    - hard optimization problems
    - neuronal networks
    - inference

what is a spin glass?

systems with frozen and disordered spin-spin interactions

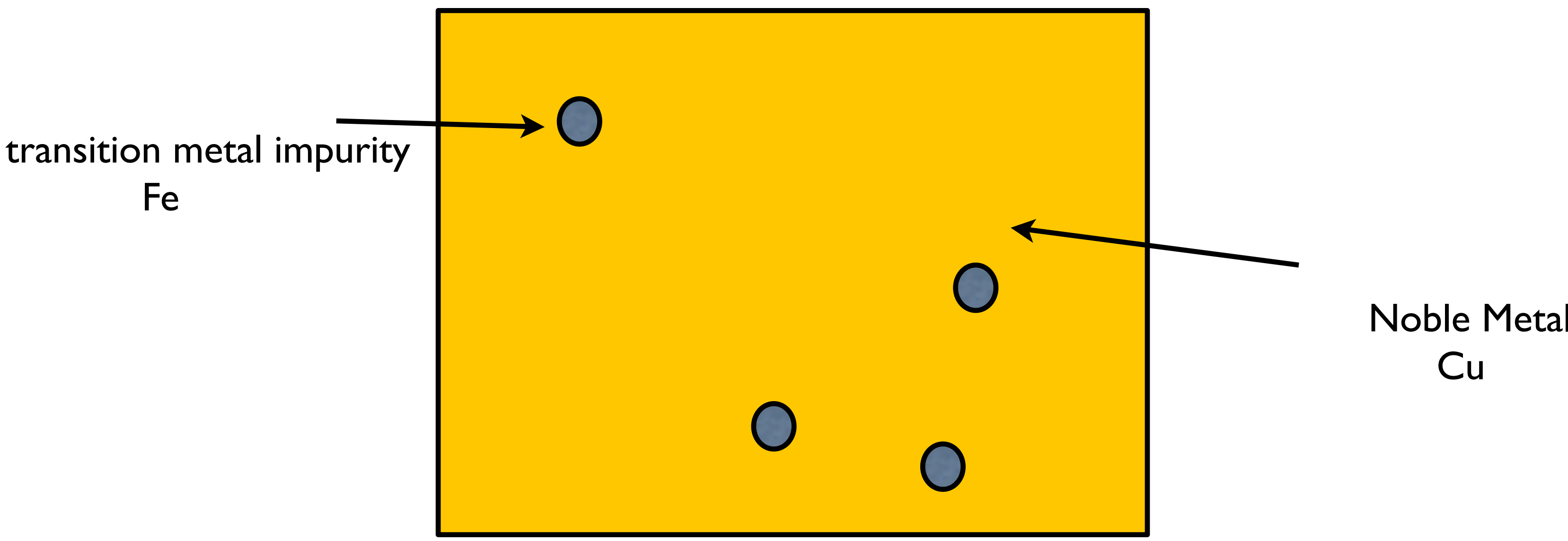
• magnetic alloys where magnetic moments interact via disordered exchange interactions.

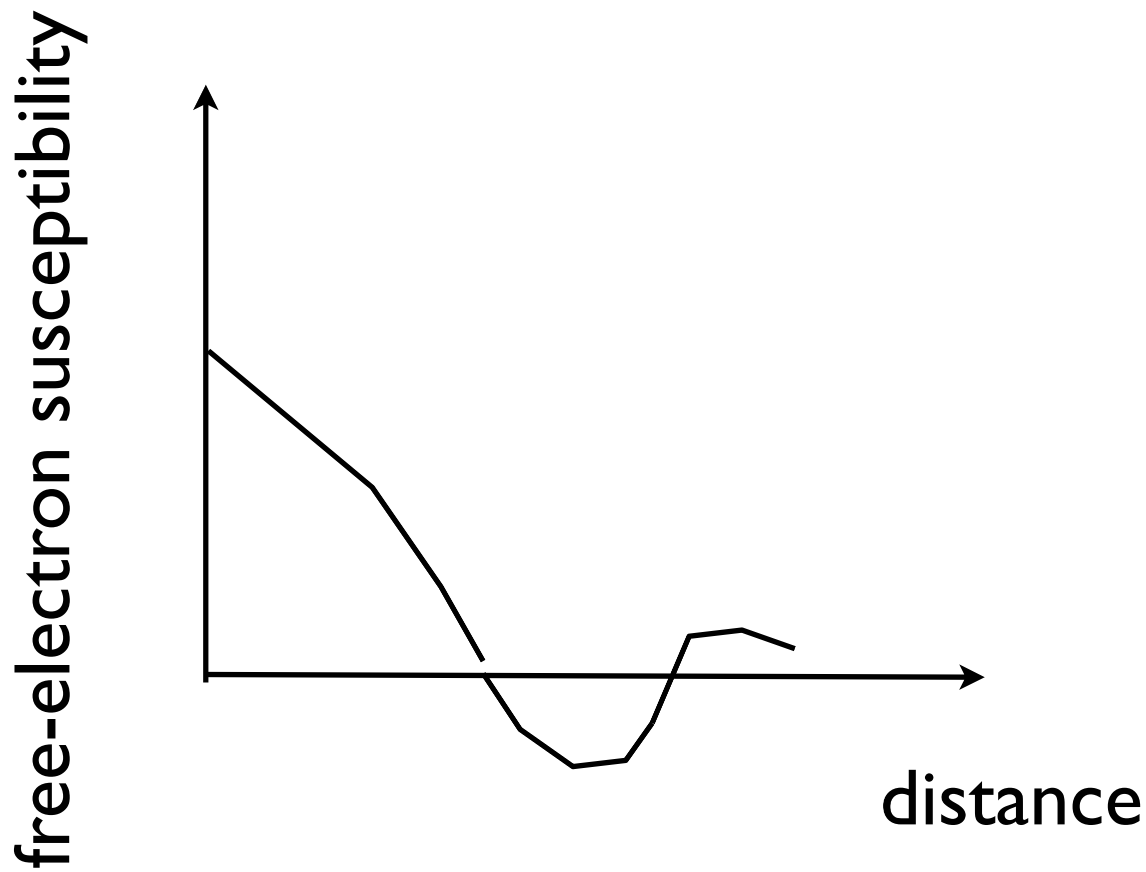




Canella and Mydosh 1974

linear temperature dependence of specific heat
time dependent remnant magnetization





$$H = -\sum_{(i,j)} J_{ij}\sigma_i\sigma_j$$

$$H = -\sum_{(i,j)} J_{ij}\sigma_i\sigma_j$$

random interactions

$$\Pr(J_{ij}) = \frac{1}{\sqrt{2\pi J^2}} \exp\left\{ -\frac{(J_{ij}-J_0)^2}{2J_1^2} \right\}$$

$$\Pr(J_{ij}) = p\delta(J_{ij}-J) + (1-p)\delta(J_{ij}+J)$$

Edwards-Anderson model: finite dimensional, short range interactions

$$m = \frac{1}{N}\sum_i \langle\sigma_i\rangle \qquad q_{EA} = \frac{1}{N}\sum_i (\langle\sigma_i\rangle)^2$$

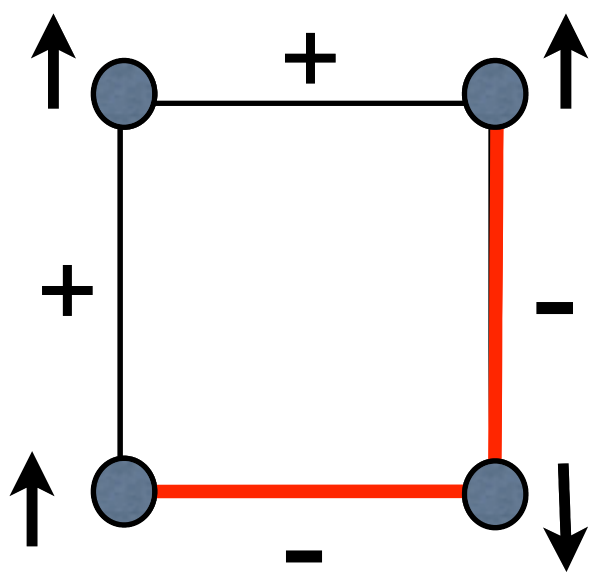
m	q_{EA}	
$= 0$	$= 0$	paramagnetic
$\neq 0$	$\neq 0$	ferromagnetic
$= 0$	$\neq 0$	frozen spins

$$\chi_{ii} = \frac{\partial m_i}{\partial h_i} = \langle(\sigma_i - \langle\sigma_i\rangle)^2\rangle = 1 - \langle\sigma_i\rangle^2 \qquad T\chi = \frac{T}{N}\sum_i \chi_{ii} = 1 - \frac{1}{N}\sum_i \langle\sigma_i\rangle^2$$

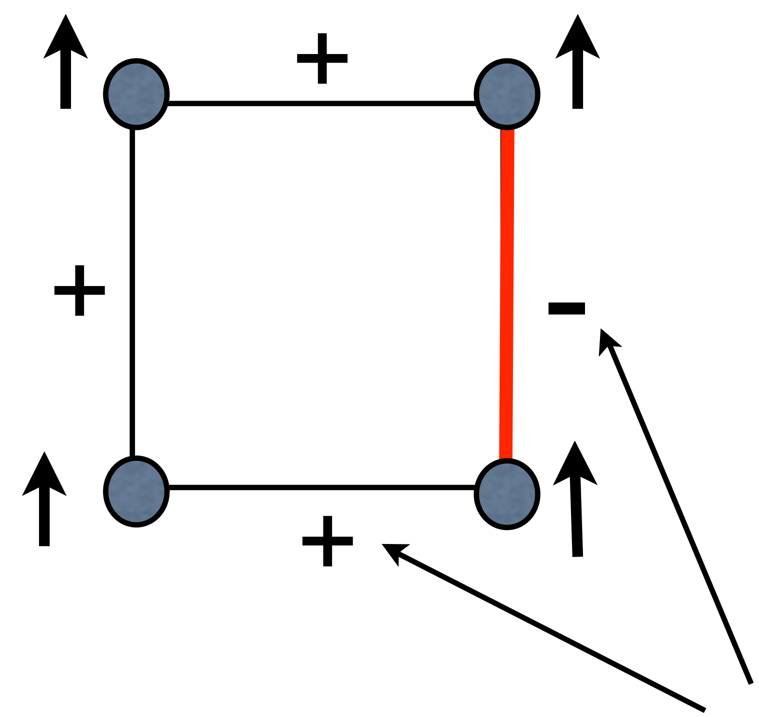
Sherrington-Kirkpatrick model: mean-field model

$$J_0 \sim \mathcal{O}(N^{-1}) \qquad J_1 \sim \mathcal{O}(N^{-1})$$

frustration



All bonds are satisfied



$$\prod_{J_{ij} \in C} J_{ij} < 0$$

how to study spin glass models?

- typical behaviour

$$\int dJ \;\; \mathrm{Pr}(J) \qquad Z[J] = \int dJ \;\; \mathrm{Pr}(J) \;\; \mathrm{Tr}_\sigma \exp\{-\beta H\}$$

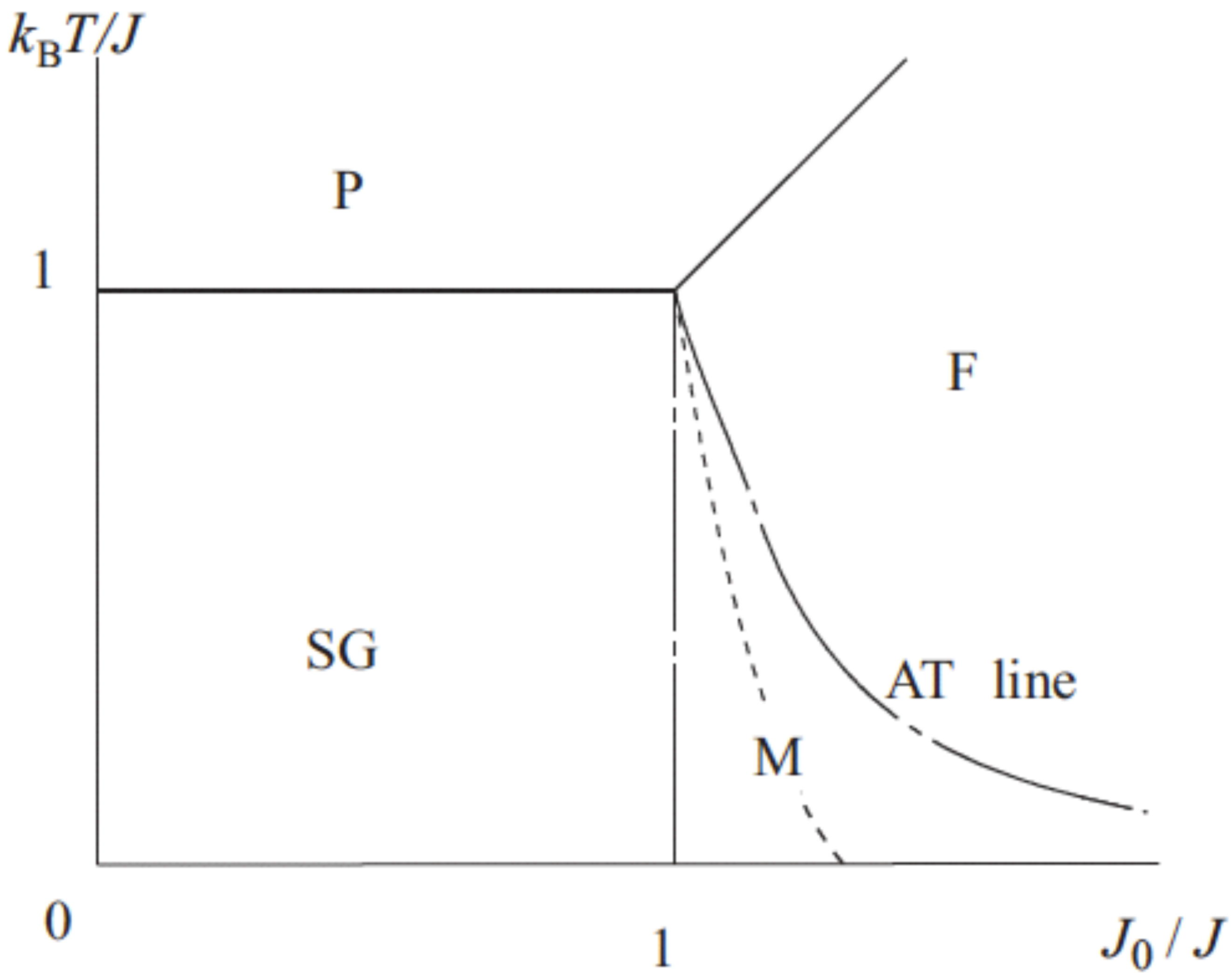
annealed approximation. Not very good!!

$$\int dJ \;\; \mathrm{Pr}(J) \;\; \log Z[J] = \int dJ \;\; \mathrm{Pr}(J) \;\; \log \mathrm{Tr}_\sigma \exp\{-\beta H\}$$

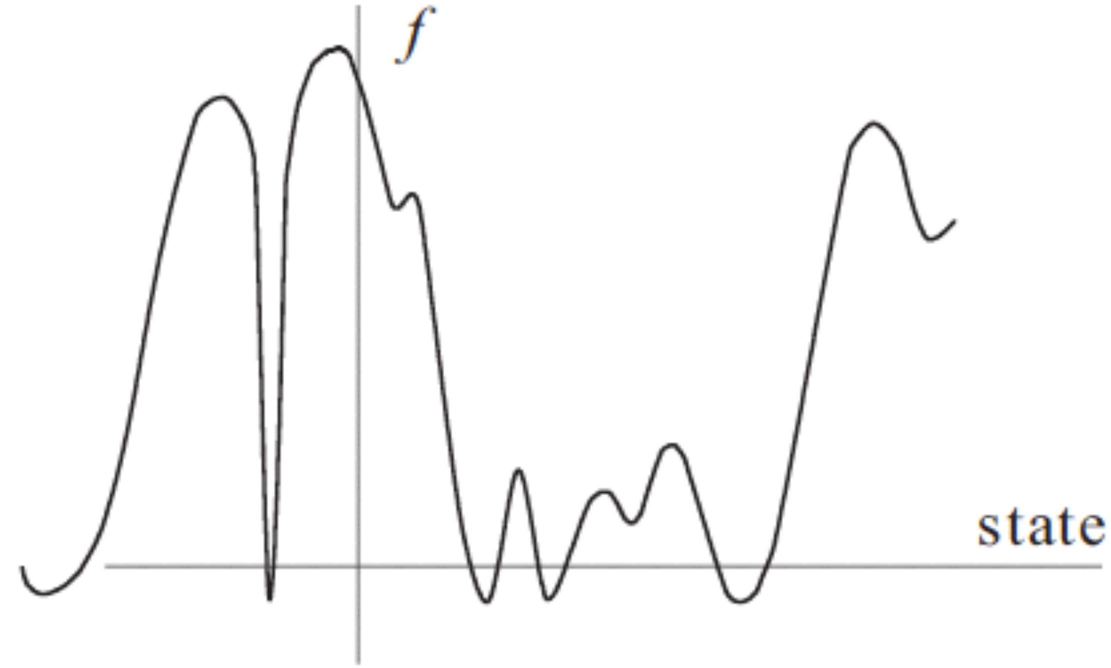
replica trick $\log Z = \lim_{n\to 0} \frac{Z^n - 1}{n}$

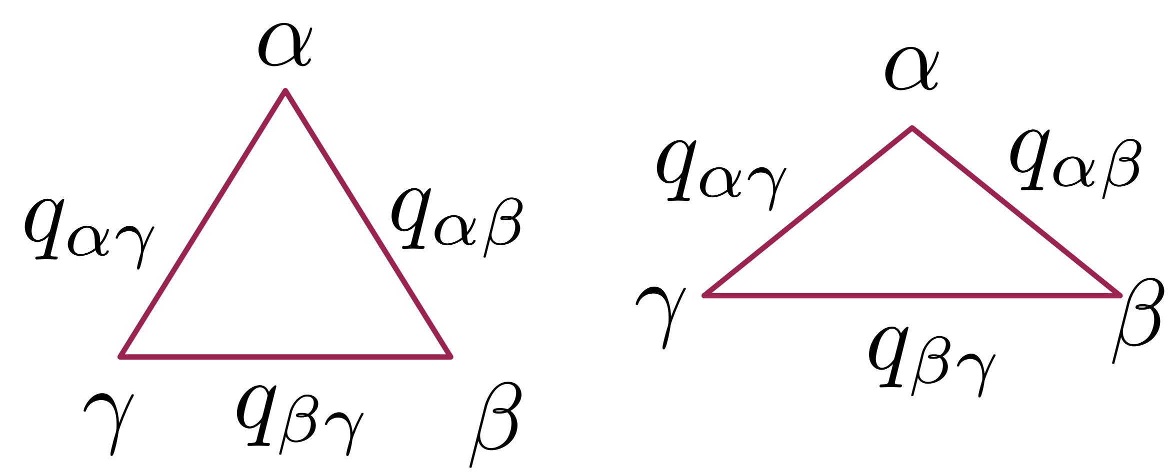
n replica of the system with the same disorder

$$q_{\alpha,\beta} = \frac{1}{N}\sum_i \langle \sigma_i^\alpha \rangle \langle \sigma_i^\beta \rangle$$

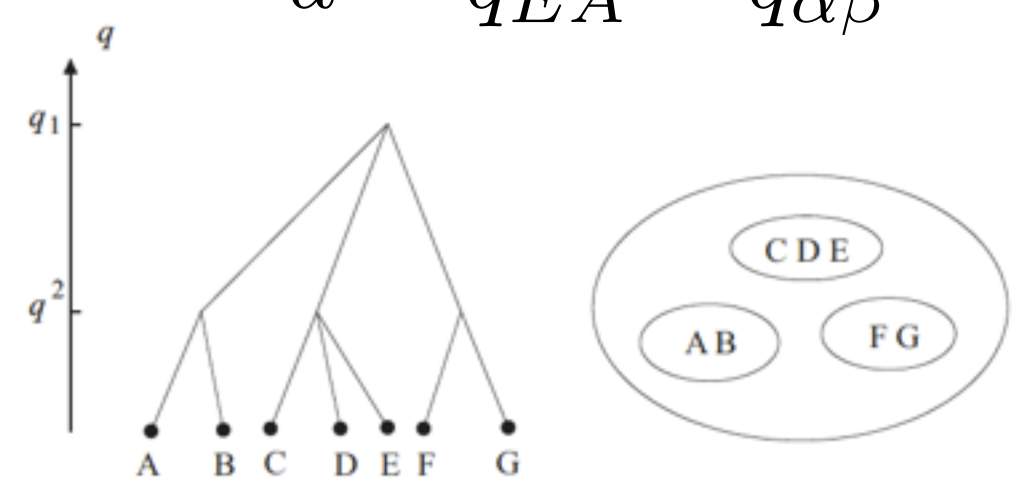


Nishimori 2001
Mezard, Parisi, Virasoro 87
Fischer and Hertz 91





$$d = q_{EA} - q_{\alpha\beta}$$



how to study spin glass models?

•behavior of a specific realization of the disorder

naive mf equations

$$m_i = \tanh[h_i + \sum_j J_{ij} m_j]$$

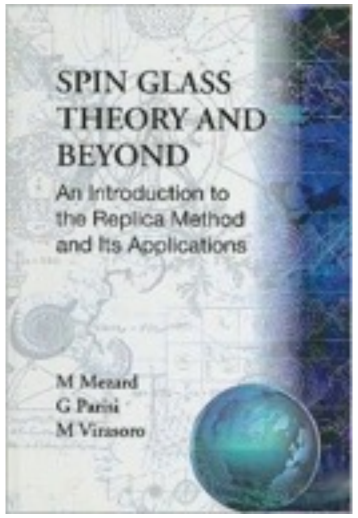
TAP equations

$$m_i = \tanh[h_i + \sum_j J_{ij} m_j - m_i \sum_j J_{ij}^2 (1 - m_j^2)]$$

formally can be derived by expanding the Gibbs free energy in powers of J (the Plefka expansion) or by one-loop corrections.

but magnetic materials are not the only systems with

random interactions and frustration



Chapter 0

A KIND OF INTRODUCTION

Often in life we find out that our goals are mutually incompatible: we have to renounce some of them and we feel frustrated. For example, I may want to be a friend of both Mr. White and Mr. Smith. Unfortunately, they hate each other: it is then rather difficult to be a good friend of both of them (a very frustrating situation).

The situation is more complex when many individuals are present. In a classical tragedy the scenario may be the following: there is a fight between two groups and the various characters on the scene have to choose sides. In addition they all have strong personal feelings, positive or negative, towards each other (it is a tragedy!). Some of them are friends and some are enemies. For simplicity we will assume that all feelings are reciprocal; otherwise the system may never reach equilibrium (this more general case, though much more complicated can be studied. See Reprint 34 for one particular example). Let us consider three characters (A, B and C); if A and B, B and C, A and C do like each other, there is no problem: they will all choose the same side. In a similar way, if A and B are friends and C is an enemy of both, then A and B can be on one side and C will be on the other. Frustration follows, instead, if A, B and C hate each other because two personal enemies must then fight on the same side.

This analysis can be formalized by assigning to each pair a number J_{AB} which is $+1$ if A and B are friends and -1 if they hate each other; the relation among three characters is frustrated if (Ref. 1)

$$J_{AB} \cdot J_{BC} \cdot J_{CA} = -1 . \tag{0.1}$$

When many triples are frustrated, evidently the situation on the scene is unstable and many rearrangements of the two fields are possible.

At a given moment of the tragedy it is possible to define the "dramatic tension" as

$$\text{Number of frustrated triples/Total number of triples}. \tag{0.2}$$

Detailed studies[2] have shown that in many Shakespeare's plays the dramatic tension has a small value at the beginning of the tragedy, reaches a maximum in the middle and decreases by the end.

Mathematically we could say that we have N variables s_i, one for each character; s_i

K-SAT problem

N boolean variables $\{x_1, x_2, \cdots, x_N\} \in \{F, T\}^N$

pick K of them x_i, x_j, x_k

construct a clause Δ_1 involving the logical OR between these variables or their negations

$$x_i \vee x_j \vee x_k$$

$$x_i \vee x_j \vee \neg x_k$$

$$x_i \vee \neg x_j \vee \neg x_k$$

do this M times

$$\Delta_1, \Delta_2, \cdots, \Delta_M$$

Questions: is there an instance of the x variables such that

$$\Delta_1 \wedge \Delta_2 \wedge \cdots \Delta_M = T$$

example

$$N = 3, K = 2, M = 3$$

$$C_1 = x_1 \vee \neg x_2$$

$$C_2 = x_1 \vee x_2$$

$$C_3 = x_2 \vee \neg x_3$$

$$x_1 = T, x_2 = T, x_3 = T$$

$$N = 3, K = 2, M = 5$$

$$C_1 = x_1 \vee \neg x_2$$

$$C_2 = x_1 \vee x_2$$

$$C_3 = x_2 \vee \neg x_3$$

$$C_4 = \neg x_1 \vee \neg x_2$$

$$C_5 = \neg x_1 \vee x_2$$

$$\alpha = M/N$$

- checking that an assignment is a solution is simple.
- finding a solution:
  - K=2, the problem is polynomial.
  - K=3, the first problem proved to be NP complete (cook 1971).

$$(x_1 \vee x_{27} \vee \bar{x}_3) \wedge (\bar{x}_{11} \vee x_3 \vee x_2) \wedge \ldots \qquad x_i = \frac{1+s_i}{2}$$

$$E = \frac{1+s_1}{2}\frac{1+s_{27}}{2}\frac{1-s_3}{2} + \frac{1-s_{11}}{2}\frac{1+s_3}{2}\frac{1+s_2}{2} + \cdots$$

$$E[\Delta,S] = \sum_{\ell=1}^{M} \delta\left[\sum_{i=1}^{N} \Delta_{\ell,i}\, S_i; -K\right]$$

$$E[\Delta,S] = \frac{\alpha}{2^K}N + \sum_{R=1}^{K} (-1)^R \sum_{i_1<i_2<\cdots<i_R} \times J_{i_1,i_2,\ldots,i_R} S_{i_1} S_{i_2} \ldots S_{i_R},$$

where the couplings are defined by

$$J_{i_1,i_2,\ldots,i_R} = \frac{1}{2^K}\sum_{\ell=1}^{M} \Delta_{\ell,i_1}\Delta_{\ell,i_2}\cdots\Delta_{\ell,i_R}.$$

VOLUME 76 20 MAY 1996 NUMBER 21

Entropy of the *K*-Satisfiability Problem

Rémi Monasson[1,*] and Riccardo Zecchina[2,†]

[1]*Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond, 75231 Paris cedex 05, France*

[2]*Istituto Nazionale di Fisica Nucleare and Dip. di Fisica, Politecnico di Torino, C.so Duca degli Abruzzi 24, I-10129 Torino, Italy*

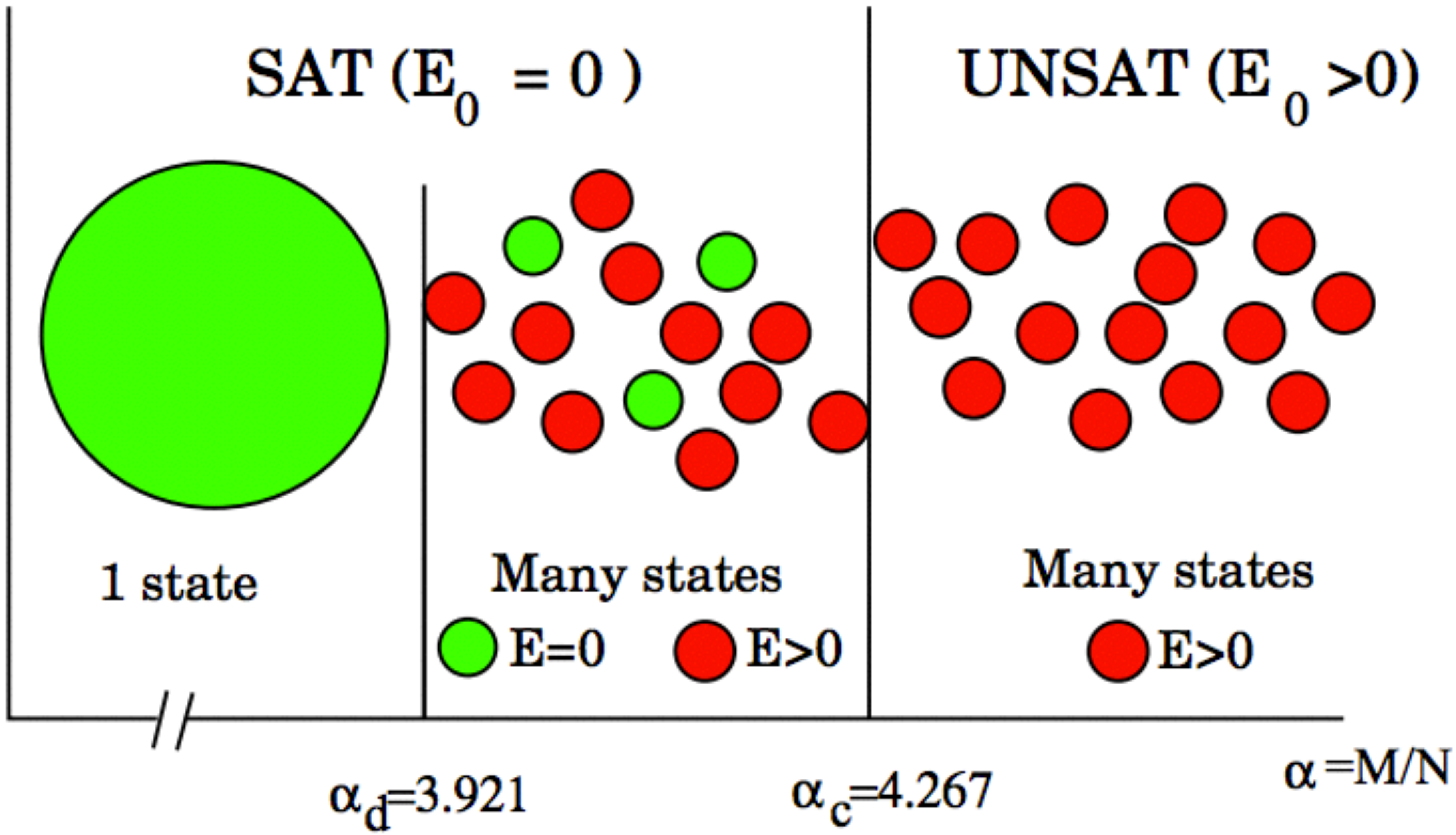
(Received 12 January 1996)

The threshold behavior of the *K*-satisfiability problem is studied in the framework of the statistical mechanics of random diluted systems. We find that at the transition the entropy is finite and hence that the transition itself is due to the abrupt appearance of logical contradictions in all solutions and not to the progressive decreasing of the number of these solutions down to zero. A physical interpretation is given for the different cases $K = 1$, $K = 2$, and $K \geq 3$. [S0031-9007(96)00244-X]

PACS numbers: 05.20.–y, 02.10.–v, 64.60.H+, 89.70.+c

K=2 smooth phase transition at $\alpha_c = 1$

K=3 sharp RS phase transition at $\alpha_c = 5.18$



- For $\alpha < \alpha_d = 3.921$, the problem is generically SAT; the solution can be found relatively easily, because the space of SAT configurations builds up a single big connected cluster. A $T = 0$ Metropolis algorithm, in which one proposes to flip a randomly chosen variable, and accepts the change iff the number of violated constraints in the new configuration is less or equal to the old one, is able to find a SAT configuration. We call this the EASY-SAT phase

- For $\alpha_d < \alpha < \alpha_c = 4.267$, the problem is still generically SAT, but now it becomes very difficult to find a solution (we call this the HARD-SAT phase).

- For $\alpha > \alpha_c$, the problem is typically UNSAT. The ground state energy density e_0 is positive. Finding a configuration with lowest energy is also very difficult because of the proliferation of metastable states.

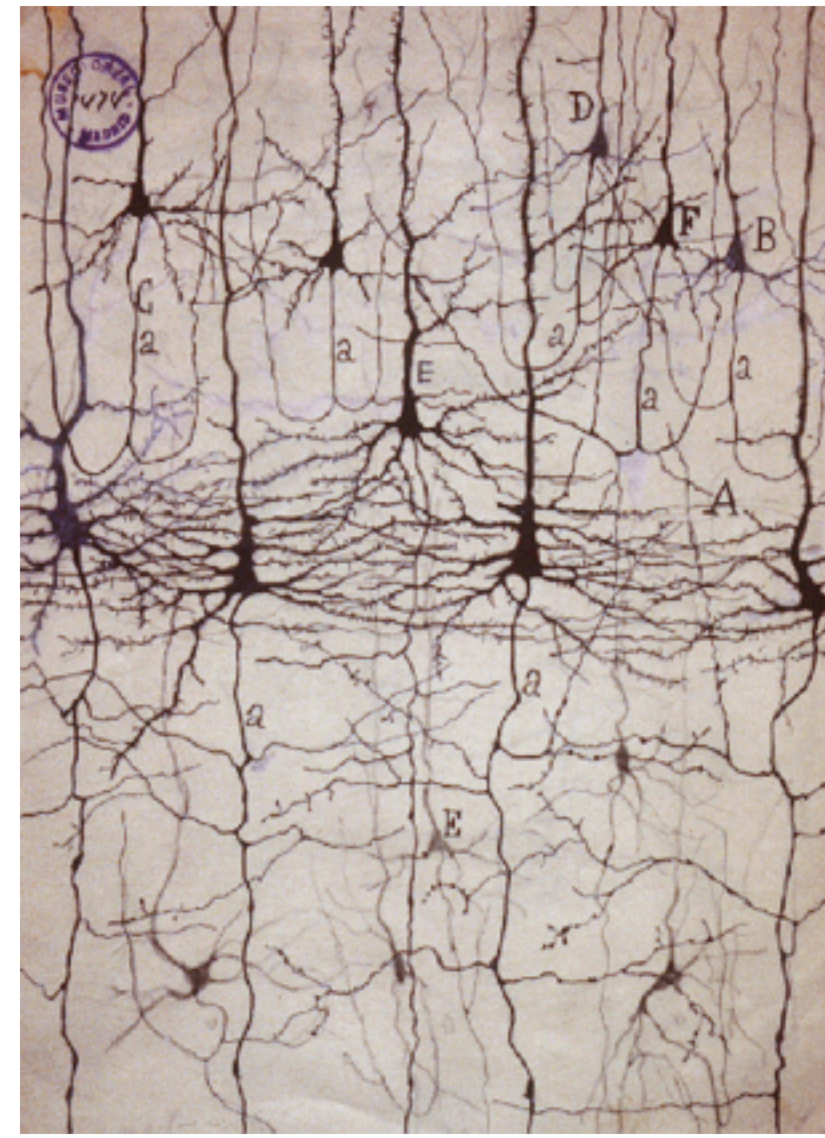
Mezard, Parisi, Zecchina, Science 2002 Mezard M, arxiv 02124448

neural networks

~10^{11} neurons in the brain

~ each connected to 10^4 others

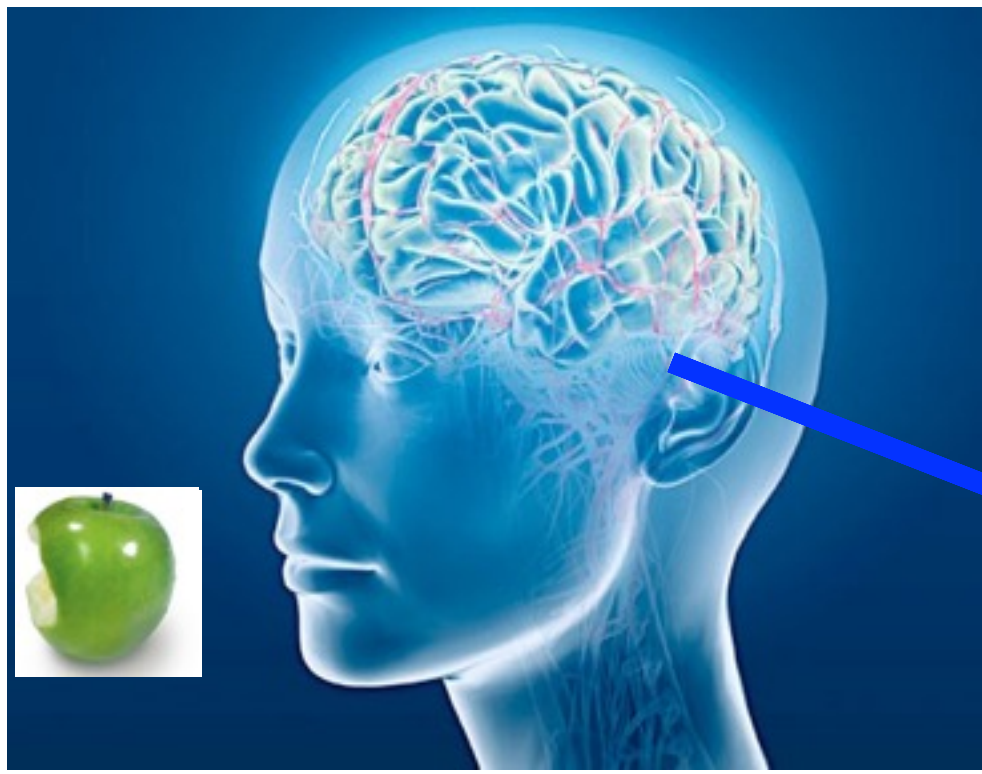
~ a typical mm^2 contains 10^5 neurons and 10^9 connections.

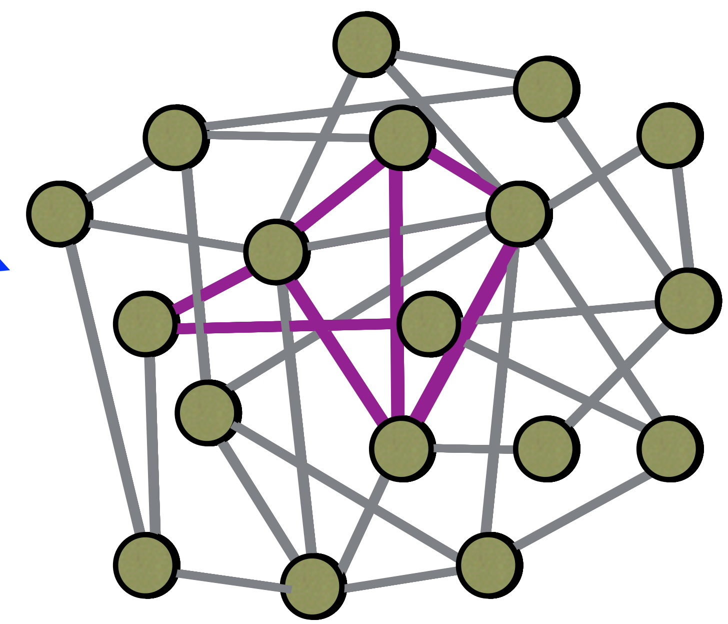


a neuron generates an action potential if it receives enough input from other neurons.

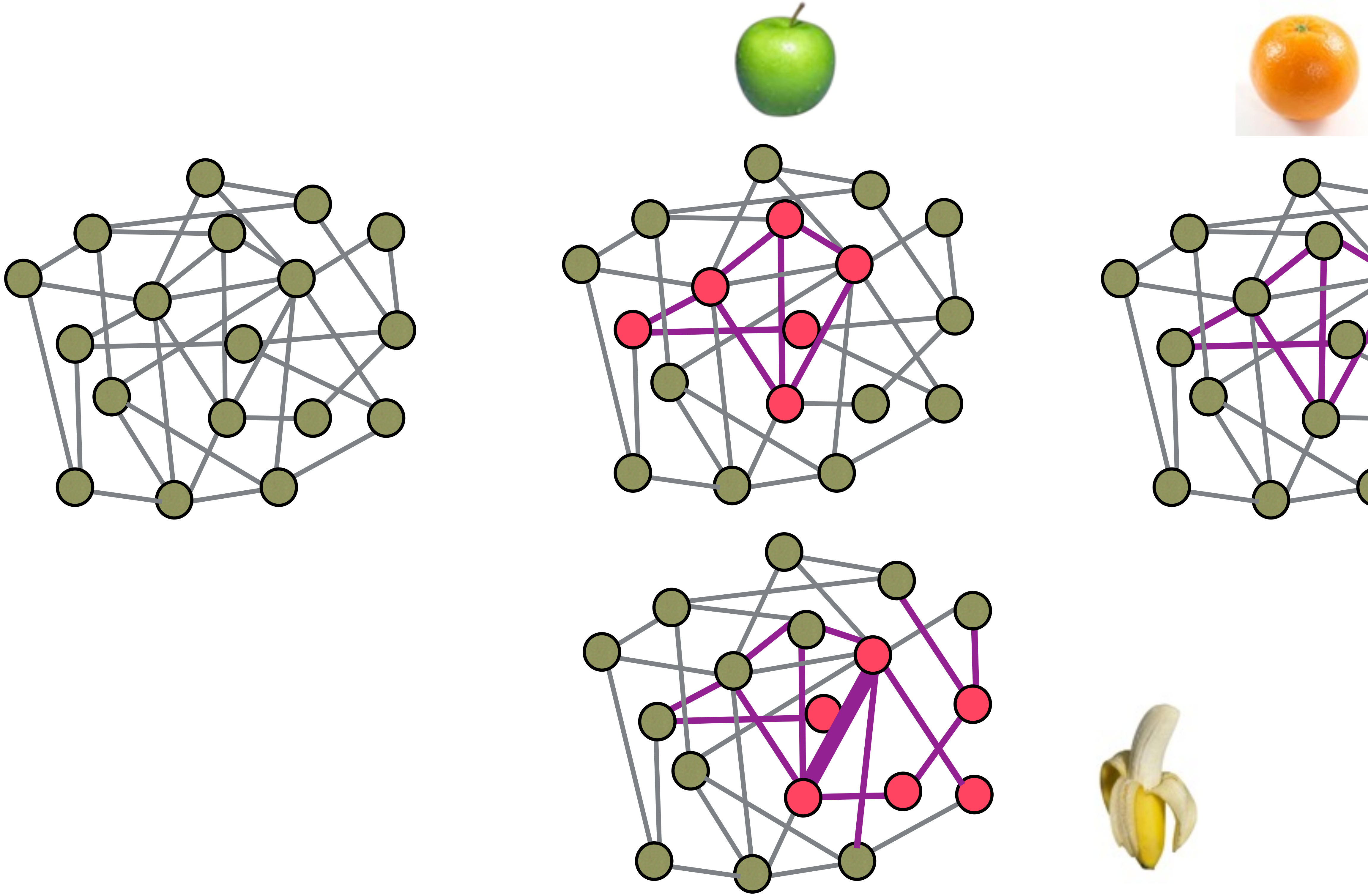
Synaptic Plasticity

the idea behind memory formation



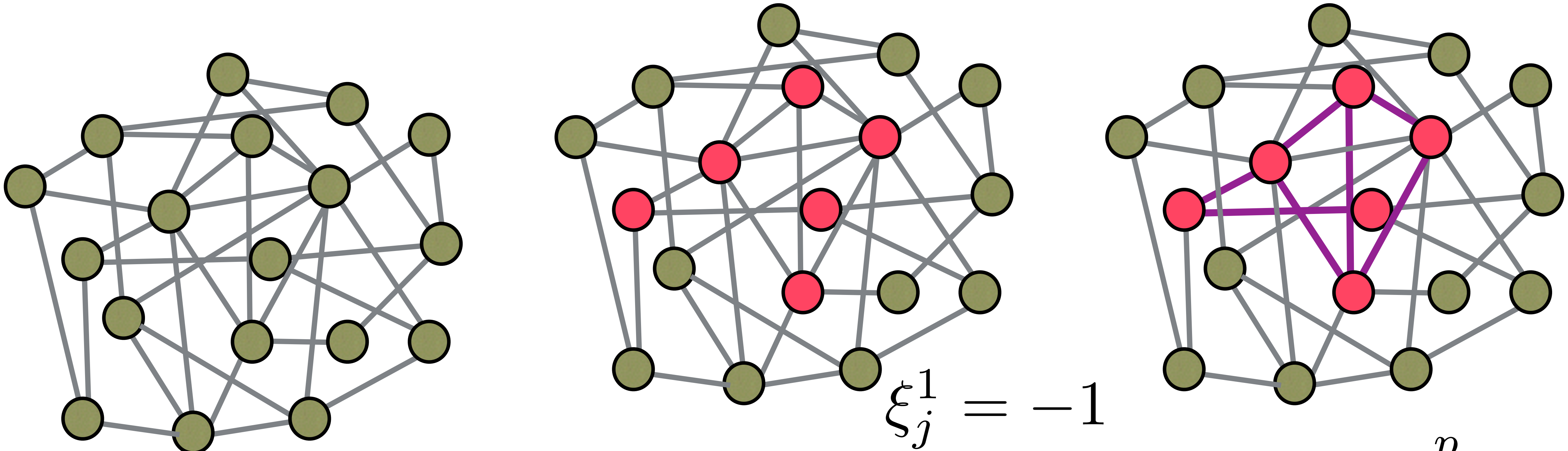


- attractor network, Hebbian cell assembly



- learning many patterns causes interference between them.

$$\xi_i^1 = 1$$



$$\xi_j^1 = -1$$

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

$$H = -\sum_{i,j} J_{ij} \sigma_i \sigma_j$$

$$F = -[\log Z]_\xi$$

$$\alpha = p/N$$

$$m^\mu = \frac{1}{N} \sum_i \xi_i^\mu \sigma_i$$

Under which conditions one (or more) of the ms will be close to 1 while the rest are close to zero?

Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks

Daniel J. Amit and Hanoch Gutfreund
Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel

and

H. Sompolinsky
Department of Physics, Bar Ilan University, Ramat Gan, Israel

(Received 11 July 1985)

The Hopfield model for a neural network is studied in the limit when the number p of stored patterns increases with the size N of the network, as $p = \alpha N$. It is shown that, despite its spin-glass features, the model exhibits associative memory for $\alpha < \alpha_c$, $\alpha_c \gtrsim 0.14$. This is a result of the existence at low temperature of $2p$ dynamically stable degenerate states, each of which is almost fully correlated with one of the patterns. These states become ground states at $\alpha < 0.05$. The phase diagram of this rich spin-glass is described.

PACS numbers: 87.30.Gy, 64.60.Cn, 75.10.Hk, 89.70.+c

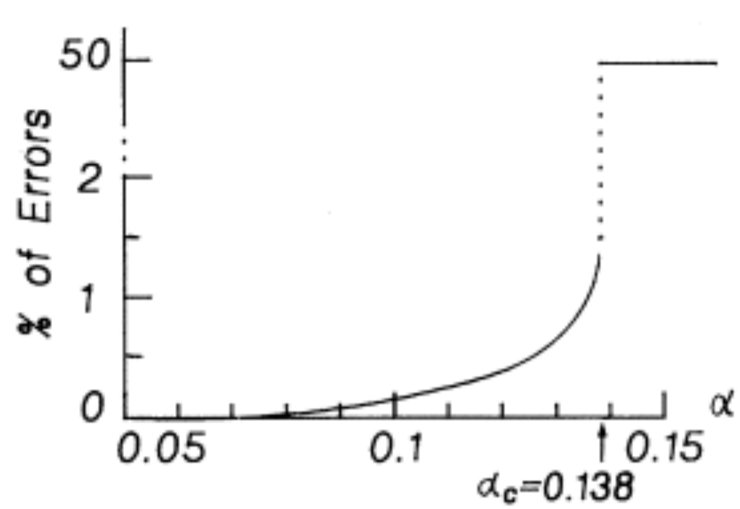
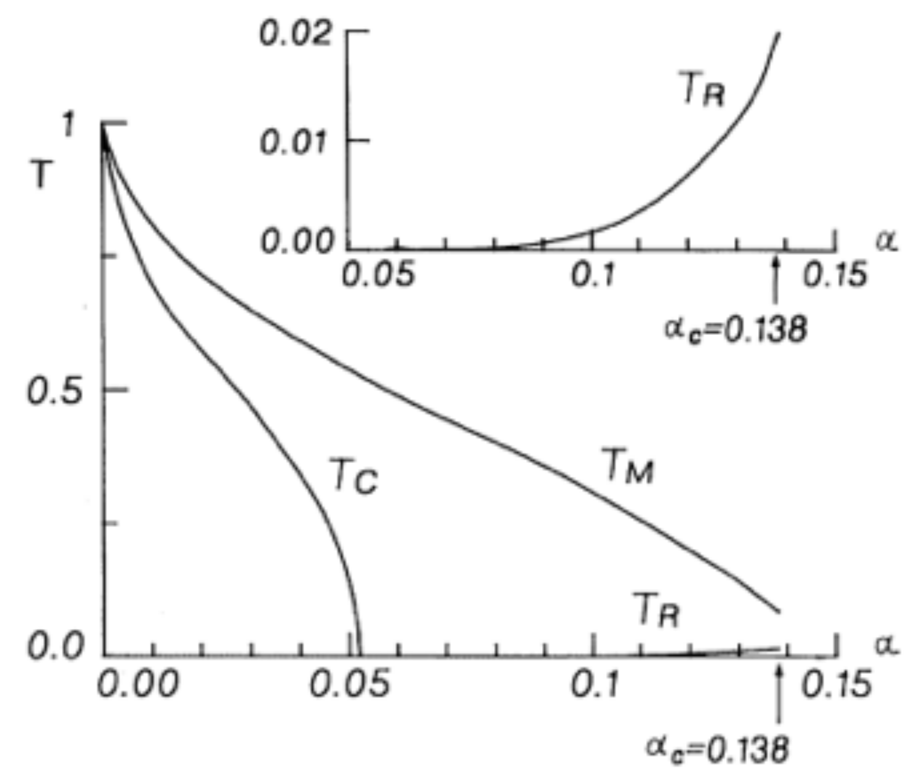
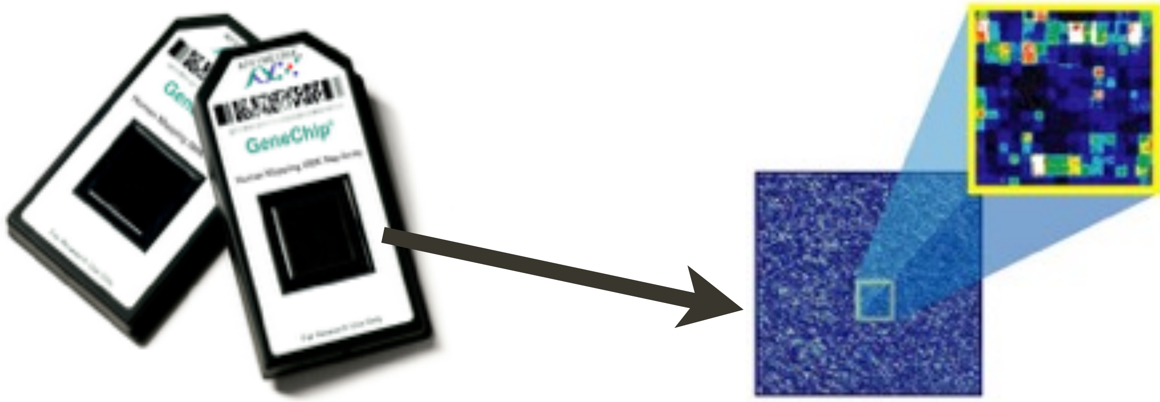
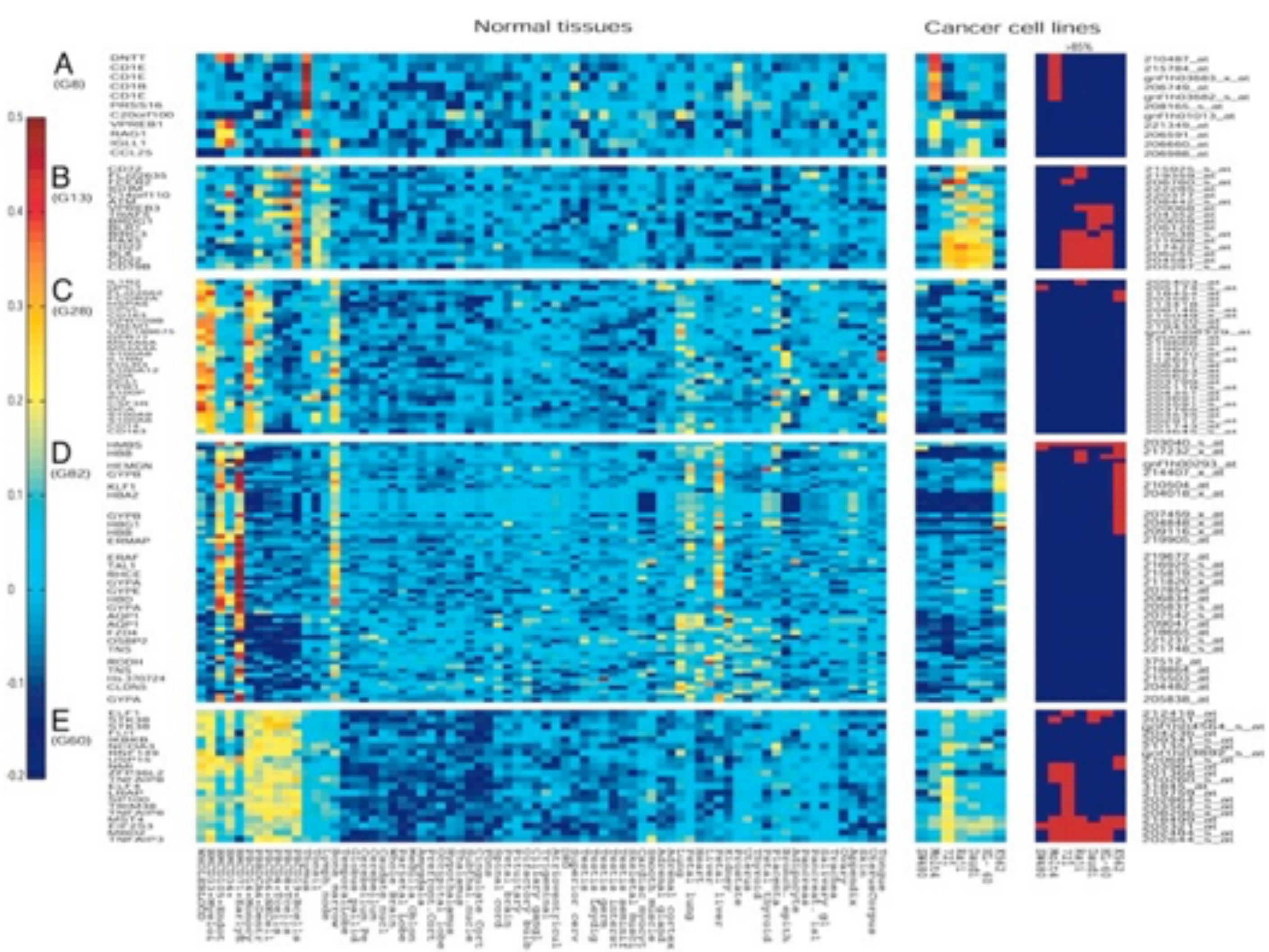


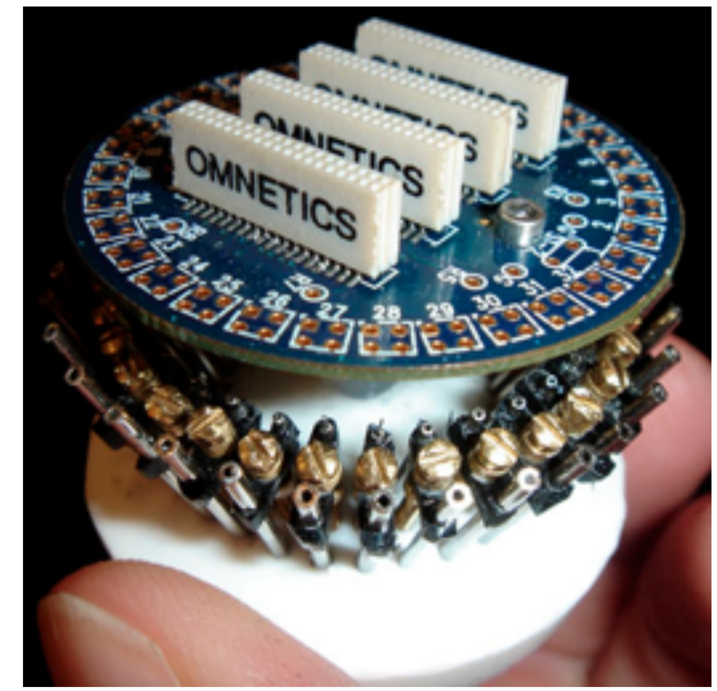
FIG. 1. Average percentage of errors in the FM states, as a function of α at $T = 0$.

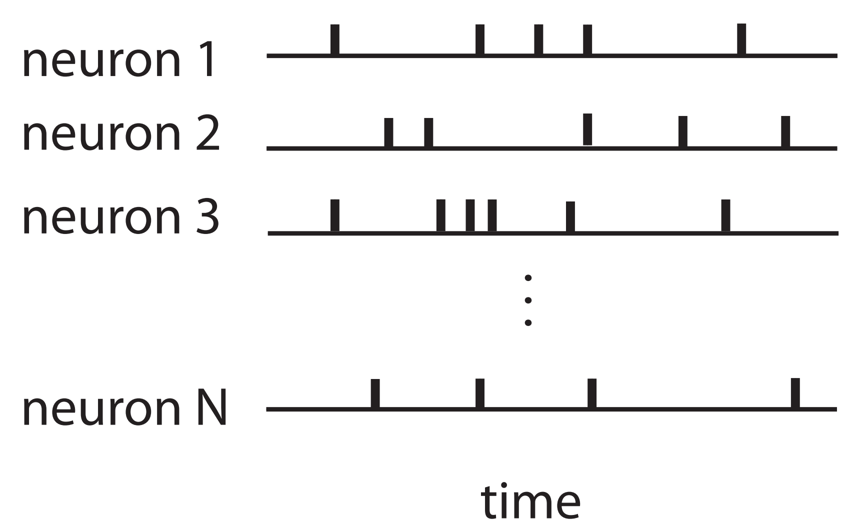


high-throughput data in biology

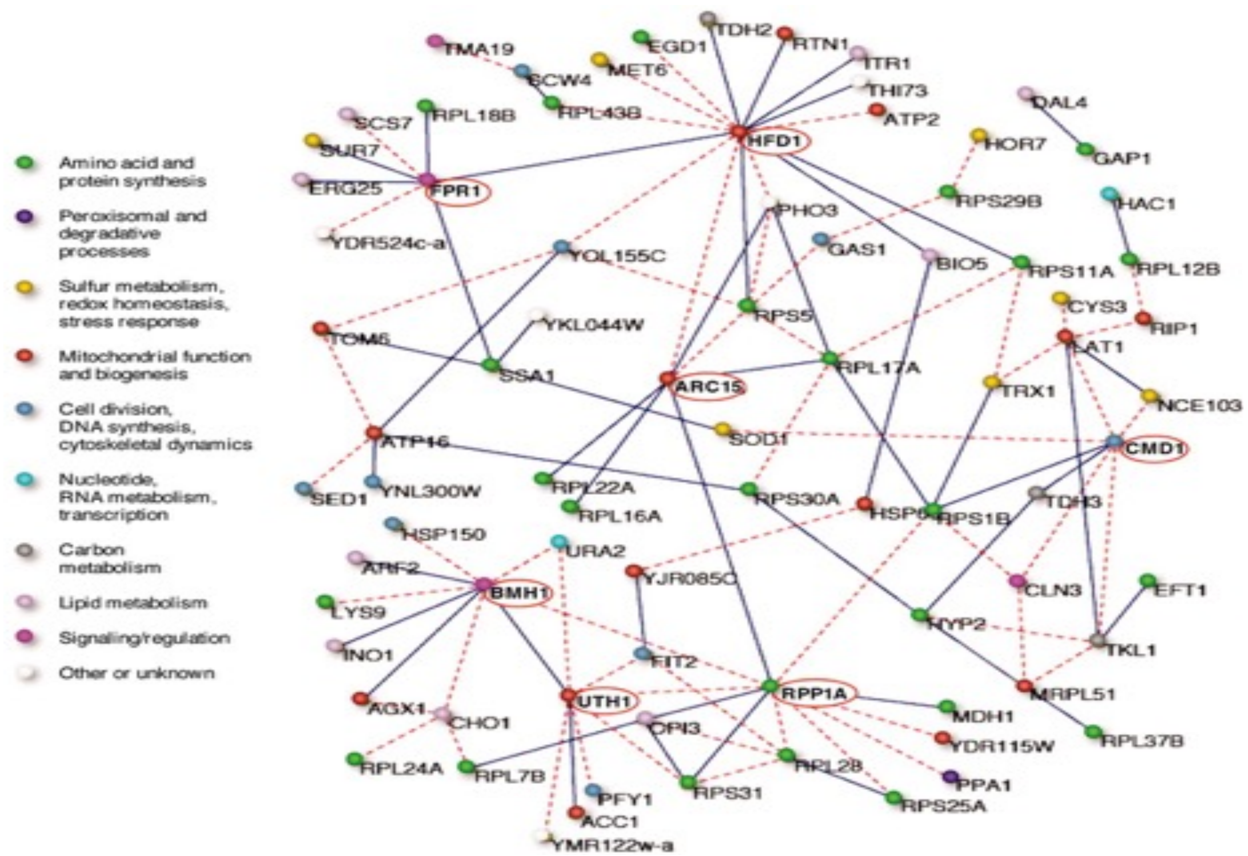








~ inferring gene regulatory network



Lazon et al 06

microarray expression data from *Saccharomyces cerevisiae*

~ Reconstructing protein complexes from co-evolution of contacting residues (#a.a. ~ 10^2; #data ~ 10^3-10^4)

Weigt, White, Szurmant, Hoch, Hwa (PNAS 2009)

equilibrium inverse Ising problem

$$m_i = \langle s_i \rangle$$

$$C_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$$

find h_i and J_{ij} of

$$\Pr(s_1, \ldots, s_N) = \frac{1}{Z} \exp \left[\sum_i h_i s_i + \sum_{i<j} J_{ij} s_i s_j \right]$$

Maximum-Likelihood approach

- the probability that the data is generated by the model at a given set of parameters (the likelihood)

- maximize the likelihood over the parameters.

- typically done iteratively

- how to find h_i, J_{ij} for large N?

Exact method: Boltzmann learning

$$\delta h_i = \eta\big[\langle s_i \rangle_{\rm data} - \langle s_i \rangle_{\rm current\ h\ and\ J}\big]$$

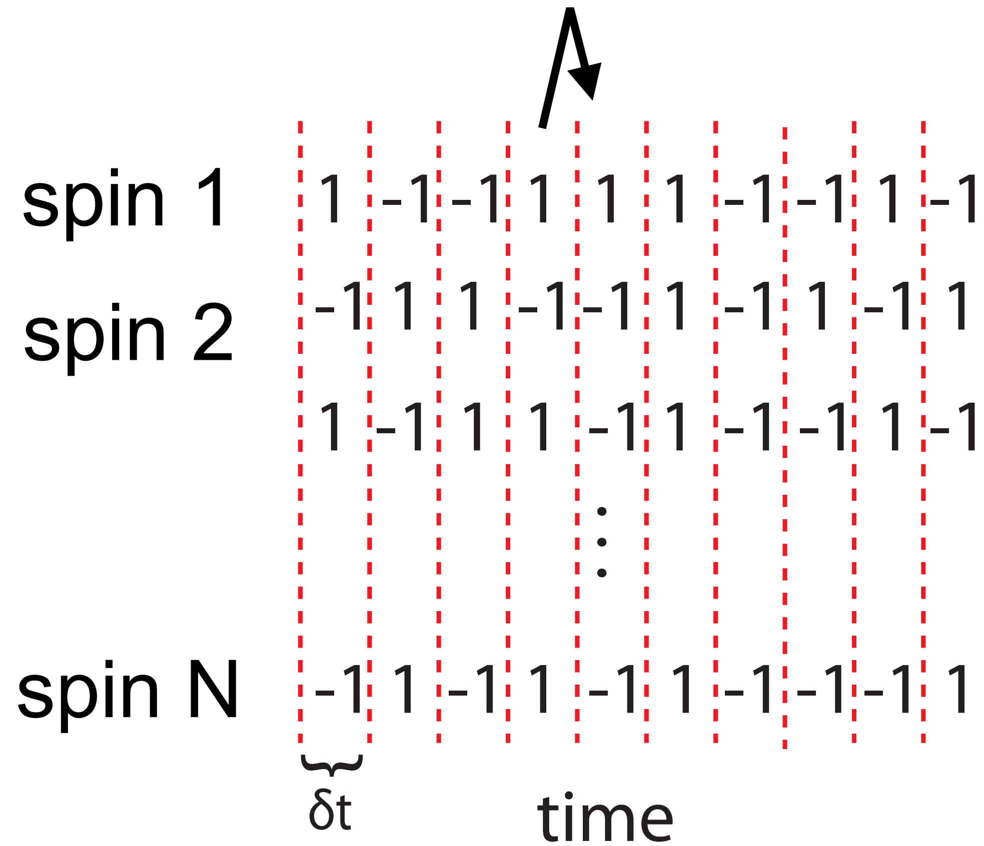
$$\delta J_{ij} = \eta\big[\langle s_i s_j \rangle_{\rm data} - \langle s_i s_j \rangle_{\rm current\ h\ and\ J}\big]$$

Ackley, Hinton, Sejnowski 85

requires long Monte Carlo runs to compute model statistics

fast and reliable approximate methods exist

inferring kinetic disordered models



$$\Pr(\{s(t+1)\}|\{s(t)\}) = \prod_i \frac{\exp[s_i(t+1)h_i(t) + \sum_j J_{ij} s_i(t+1) s_j(t)]}{2\cosh[h_i(t) + \sum_j J_{ij} s_j(t)]}$$

J_{ij} is not necessarily symmetric and the system may never reach equilibrium

- inverse problem: suppose we have observed R repeats each of length L of the spin history

$$\boldsymbol{s}^r(t) = \{s_1^r(t), \cdots, s_N^r(t)\},\ r = 1 \ldots R.$$

we would like to find out the couplings J and the fields h

J and h $\longleftrightarrow$ spin history
correlation functions

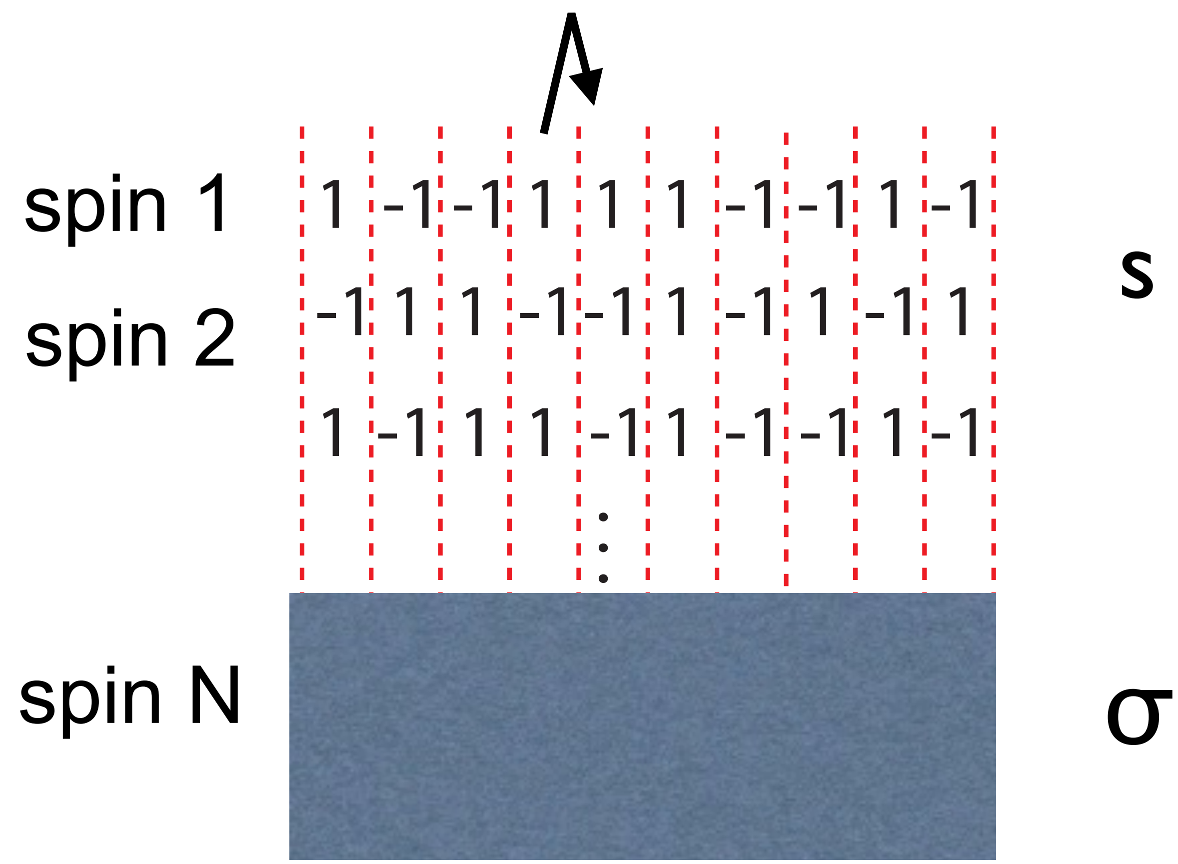
$$\delta h_i(t) = \eta_h \{ \langle s_i(t+1) \rangle_r - \langle \tanh[h_i(t) + \sum_k J_{ik} s_k(t))] \rangle_r] \}$$

$$\delta J_{ij} = \eta_J \{ \langle s_i(t+1) s_j(t) \rangle - \langle \tanh[h_i(t) + \sum_k J_{ik} s_k(t)] s_j(t) \rangle \}$$

like (batch version) delta-rule for N independent perceptrons

Much faster than Boltzmann learning for the symmetric case because it doesn't need long Monte Carlo runs to evaluate the second term

what if we don't see all spins?



$$\mathrm{p}(\{s,\sigma\}(t+1)|\{s,\sigma\}(t)) = \exp\left[\sum_i s_i(t+1)g_i(t) + \sum_a \sigma_a(t+1)g_a(t)\right]Z(t)^{-1}$$

$$Z(t) = \prod_{i,a} 2\cosh[g_i(t)]\ 2\cosh[g_a(t)]$$

$$g_i(t) = \sum_j J_{ij}s_j(t) + \sum_b J_{ib}\sigma_b(t)$$

$$g_a(t) = \sum_j J_{aj}s_j(t) + \sum_b J_{ab}\sigma_b(t)$$

Dunn and Roudi, PRE, 2013

- learn the Js: we need to calculate the likelihood of the data

$$p[\{s(t)\}_{t=1}^{T}] = \mathrm{Tr}_\sigma \prod_t p[\{s,\sigma\}(t+1)|\{s,\sigma\}(t)]$$

- infer the state of hidden spins: we need the posterior

$$p[\{\sigma(t)\}_{t=1}^{T}|\{s(t)\}_{t=1}^{T}] = \frac{p[\{\sigma(t)\}_{t=1}^{T}\ , \{s(t)\}_{t=1}^{T}]}{p[\{s(t)\}_{t=1}^{T}]}$$

if we calculate,

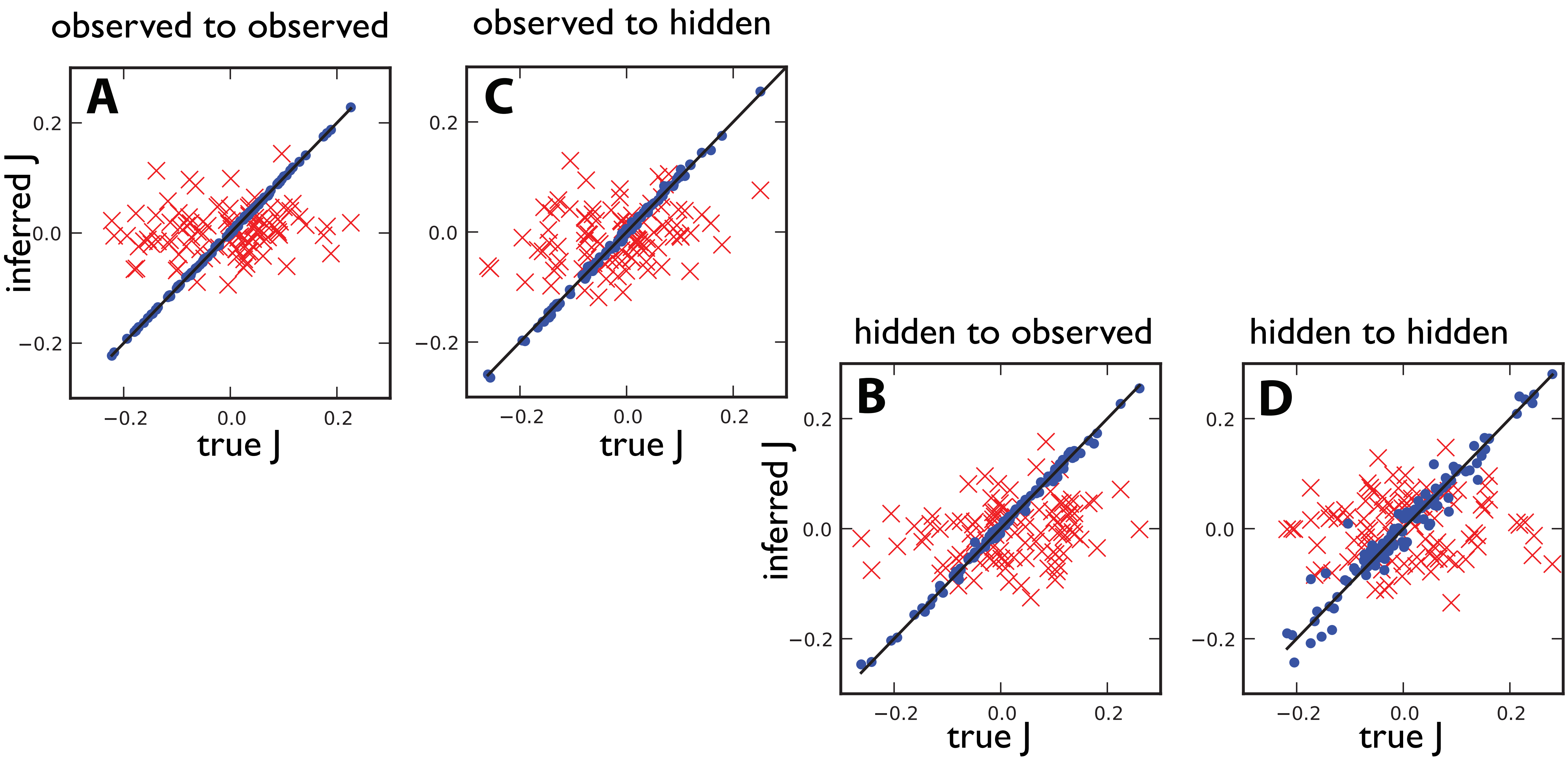
$$\mathcal{L}[\psi] \equiv \log \mathrm{Trace}_\sigma \prod_t e^{\sum_a \psi_a(t)\sigma_a(t)} p[\{s,\sigma\}(t+1)|\{s,\sigma\}(t)].$$

we have both the likelihood and the posterior

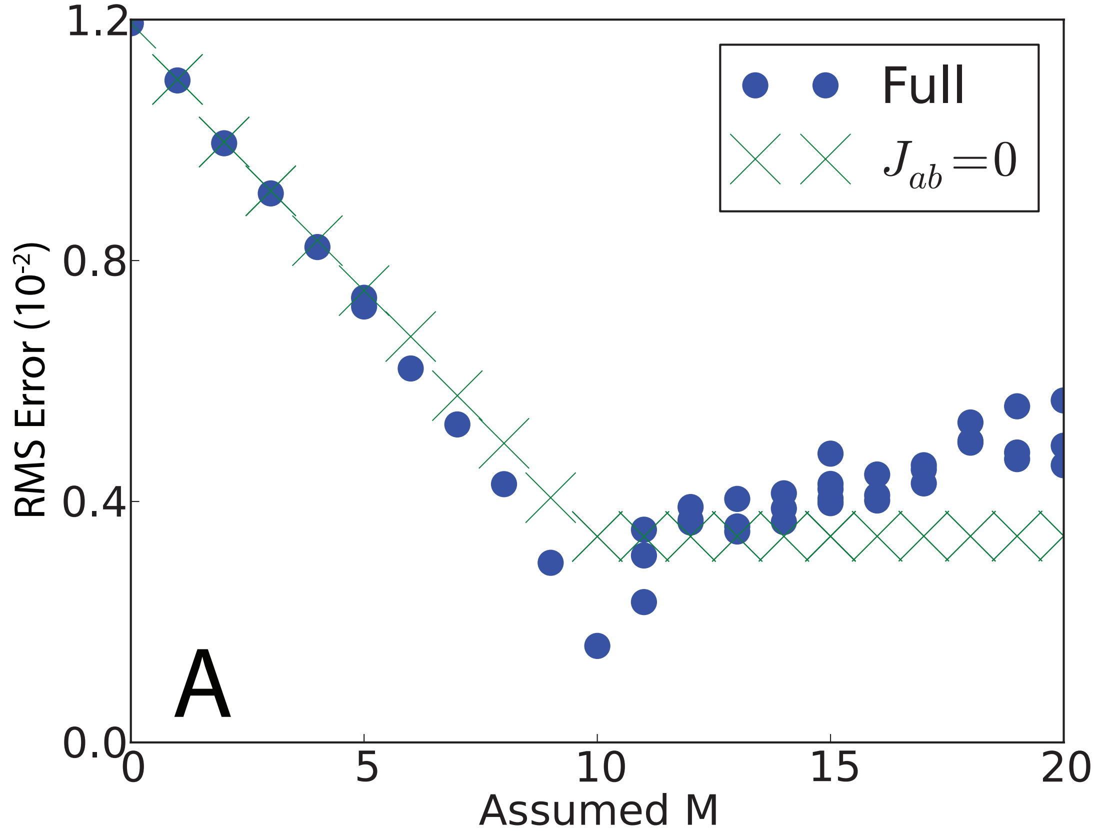
but this involves a trace over $\{\sigma(1), \ldots, \sigma(T)\}$

$$\mathcal{L}[\psi] \equiv \log \mathrm{Trace}_\sigma \prod_t e^{\sum_a \psi_a(t)\sigma_a(t)} p[\{s,\sigma\}(t+1)|\{s,\sigma\}(t)].$$

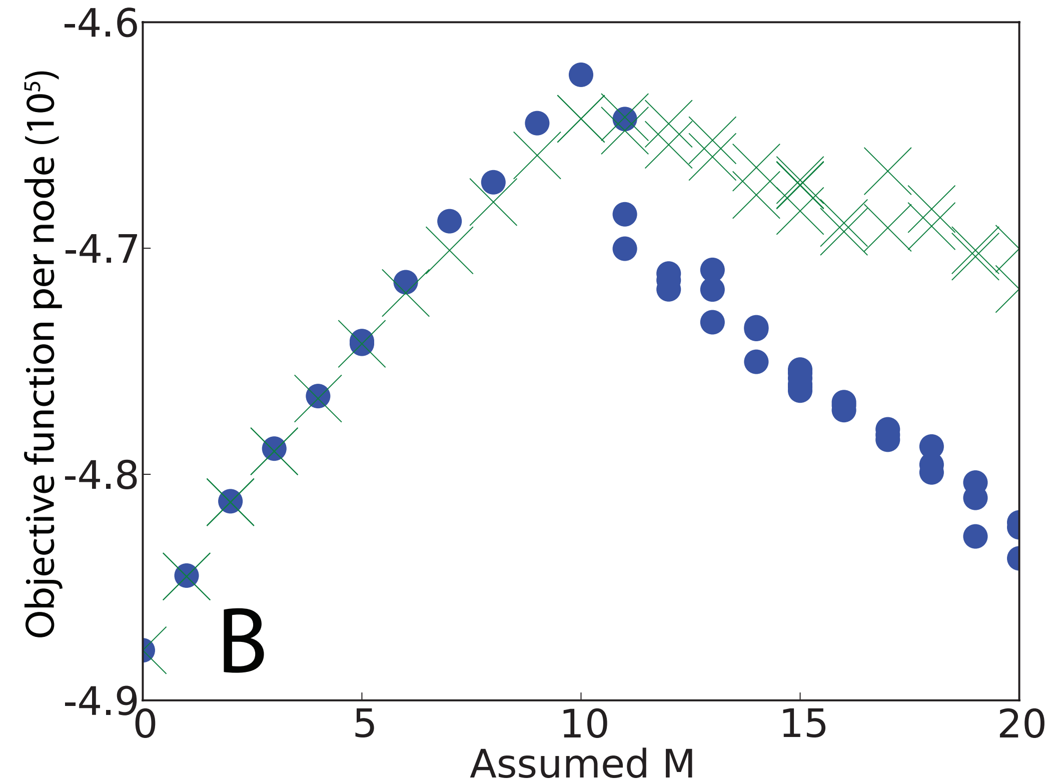
This can be written as a path integral and studied using a Plefka-like expansion.



what if we did not even know that there are hidden spins?



$$\frac{L}{N_{observed} + M}$$



summary

- spin glasses show interesting physics and analyzing them yields powerful tools.

- many other complex systems show similar behaviors, and can be analyzed using the same tools including optimization problems, learning in neural networks, statistical inference.

Norwegian Centre of Excellence

THE KAVLI FOUNDATION

www.ntnu.no/kavli

NETADIS

Statistical Physics Approaches to Networks Across Disciplines

www.netadis.eu